International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

ENHANCING SECURITY USING KEYSTROKE DYNAMICS

Preetha S
Department of ISE
BMSCE, Bangalore,
Karnataka, India

Suhas Kini K
Department of ISE
BMSCE, Bangalore,
Karnataka, India

Shailesh A Patil
Department of ISE
BMSCE, Bangalore,
Karnataka, India

Abstract—T**he uniqueness and usefulness of keystroke dynamics is an authentication measure using the typing rhythm of the user. A measure to incorporate keystroke dynamics in daily authentication systems has not been made. User's identity is analyzed using keystroke through their way of typing on a computer keyboard. Keystroke is a genuinely software centric solution and can be used for long samples of text which further can be used to increase the robustness of the method. Keystroke dynamics is an underappreciated system and can be used extensively as a robust authentication tool either individually or as a hybrid system. Knowing the cost effectiveness and efficiency we discovered that a much better use of keystroke analysis for authentication needs to be used. We also introduce keystroke dynamics based authentication and present distance metric or scoring based machine learning models. The proposed work develops a Key logger to record the keystrokes of the user and stores it. A custom built feature extractor for extracting features from key logger data is developed and also incorporates different classification models to classify and measure the similarity. Finally authentication of the user is done based on the results.**

Keywords— **keystroke dynamics; authentication; security; biometric.**

I. INTRODUCTION

Cyber security has become a major issue in recent years. The efficiency of hackers are slowly making traditional password protection obsolete and leading to a new era of biometric security. Keystroke analyses the user's identity through their way of typing on a computer keyboard. The keystroke behavior of users are unique and thus cannot be imitated, stolen, forgotten or misplaced. It is based on the assertion that a person's keystroke behavior is exactly as unique as their fingerprint. The traditional password procedure on the internet can be completely replaced by keystroke analysis.

Biometric Security systems such as fingerprint sensors, facial recognition cameras have their own constraints which makes it improbable to install and implement everywhere easily. Facial recognition might not work in dark places. Fingerprint might not work if the finger in question is damaged or dirtied and so on. All these constraints show the need for a cheaper and effective biometric security measure. Keystroke dynamics works as the perfect solution that is both effective and affordable.

Keystroke behavior of users is distinctive. The assertion that a person's keystroke behavior is exactly as unique as their fingerprint was borne out by scientific research at the University of Regensburg. The scope of keystroke analysis is illustrated in the Figure 1. Intrusion detection system has an accuracy of limiting the False alarms. The system is very robust and will not authenticate intruders even if they use typing rhythms of a language that is different. Hence keystroke analysis is the next best solution to authenticate personal identity easily.

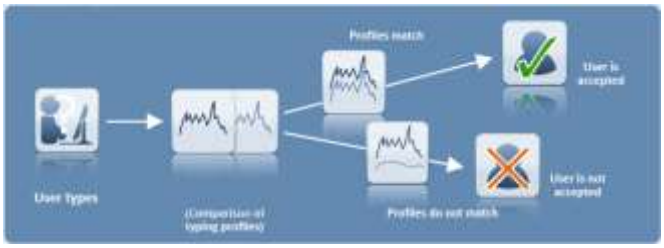

Fig. 1 Keystroke Analysis

Previous modes of authentication systems have their own constraints and limitations. The drawbacks in case of password are

- The password has to be changed regularly (Password Aging).
- Old passwords should not be re-used after a password change (Password History).
- Trivial passwords and easy-to-guess words like names or car code plates should be prevented.
- After a small number of wrong inputs the access has to be blocked for at least a limited time (Intruder Lockout).

Fingerprint recognition: Finger prints are highly unique for every individuals and it is said that "No two fingerprints are alike". In recent times fingerprint authentication is the most

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

widely accepted technology used. An individual's fingerprint contains a myriad of ridges and valleys along with minutiae points. Minutiae points define the characteristics of local ridges that show the bifurcation of an ending. The three methods for scanning finger prints are Optical scanners, Thermal scanners and Capacitance (solid state) scanners. The two accepted methods for extracting the fingerprint are Minutia-based and Correlation-based.

Minutia-based is the more microscopic of the two. This method locates the ridge characteristics (branches and endings) and assigns them a XY-coordinate that is then stored in a file. The correlation-based method looks at the entire pattern of ridges and valleys in the fingerprint. The location of the whorls, loops and arches and the direction that they flow in are extracted and stored. Both the method stores only data discarding the captured image; making it impossible to recreate the fingerprints. After the completion of scanning, a comparison of the aforementioned minutia is done and analyzed. Investigators are systems that look at, where the ridge lines end or where one ridge splits into two (bifurcation).

The scanning system uses complicated algorithms to recognize and analyze the minutia. If two prints have three ridge endings, two bifurcations, and form the same shape with the same dimensions, then it is likely the same person's fingerprints. A template as such cannot be used to recreate finger images and people in rural areas consider fingerprint authentication as associated with criminal activity are some of the drawbacks of such systems.

Facial recognition: Facial recognition is a comparatively newly arrived technology. It consists of a digitized database of pictures to which the facial characteristics of the user will be matched. Facial recognition technology is relatively new and has only been commercially available since 1990's. Face recognition has received a surge of attention since disaster of 11/9 for its ability to identify known terrorists and criminals.

Face recognition uses distinctive features of the face – including the upper outlines of the eye socket, the areas surrounding the cheekbones, the sides of the mouth, and the location of the nose and ears – to perform verification and identification. Firstly, the template image of the user is obtained and stored. The software might ask individuals to take a series of pictures with different expressions as well. Next, the images are analyzed and extracted to create a template. The last step is to verify the individual's identity by matching images to those images that been stored in database.

The four main methods being used for facial recognition are

- *Eigen faces:* It uses 2D grayscale imagery to extract features and was developed by Massachusetts Institute of Technology.
- *Feature Analysis:* One of the most widely used technique because of its ability to accommodate for facial changes and aspect. LFA uses an algorithm to create a face print (84 bytes in size) for comparison.
- *Neural network:* The process that helps creates a template image using several pictures and then helps match various elements of the template.
- *Automated Face Processing (AFP):* The distances between different elements of the facial feature are measured and hence are more suitable for low light conditions.

Aging of an individual may confuse the system, overt exposure to light may hamper the performance and limited efficiency and performance are some of the drawbacks of facial recognition algorithms.

Iris recognition: Iris comprises of over 500 unique characteristics. Comparatively iris compares more features than a fingerprint authentication system. Therefore, iris scanning is much more accurate than fingerprints or even DNA analysis of the distinguishing features. In identifying one's Iris, Iris identification systems use Passive and Active methods. The active Iris system method requires that a user be anywhere from six to 14 inches away from the camera. It also requires the user to move back and forth so that the camera can adjust and focus in on the user's iris. The passive system allows the user to be anywhere from one to three feet away from the camera(s) that locate and focus in on the iris. The technology's main uses are for authentication, identification, and verification of an individual. Instructions to safely use and operate optical readers in an Iris recognition system is cumbersome to learn. Also Iris recognition systems are intrusive in nature.

II. LITERATURE SURVEY

An authentication system that applies machine learning techniques to observe a user's cognitive typing rhythm was proposed by Chang and J. Morris (2013) in [1]. Results from a large-scale experiment at Iowa State University show the system's effectiveness. The work focuses on using behavioral biometrics, extracted from keystroke dynamics, as something a user is for active authentication. This scheme performs continual verification in the background, requires no additional hardware devices and is invisible to users. Vector representations for language have been shown to be useful in a number of Natural Language Processing tasks. Xian Fan et.al (2016) in [2] aimed to investigate the effectiveness of word vector representations for the problem of Sentiment Analysis. Particularly they target three sub-tasks namely sentiment

words extraction, polarity of sentiment words detection, and text sentiment prediction. Effectiveness of vector representations over different text data and evaluate the quality of domain-dependent vectors was investigated. Vector representations have been used to compute various vector-based features and conduct systematically experiments to demonstrate their effectiveness.

Due to the increasing vulnerabilities in cyberspace, security alone is not enough to prevent a breach, but cyber forensics or cyber intelligence is also required to prevent future attacks or to identify the potential attacker. The unobtrusive and covert nature of biometric data collection of keystroke dynamics has a high potential for use in cyber forensics or cyber intelligence. A study of the usefulness of keystroke dynamics to establish the person identity was done by Mondal, Soumik, and Patrick Bours (2017) in [3]. They proposed three schemes for identifying a person when typing on a keyboard. They used various machine learning algorithms in combination with the proposed pairwise user coupling technique and show the performance of each separate technique as well as the performance when combining two or more together. Pairwise user coupling in a bottom-up tree structure scheme gives the best performance, both concerning accuracy and time complexity. The proposed techniques are validated by using keystroke data. The ability to recognize emotions is an important part of building intelligent computers. Emotionally-aware systems would have a rich context from which to make appropriate decisions about how to interact with the user or adapt their system response. The problems with current system approaches for identifying emotions that limit their applicability is that, they can be invasive and require costly equipment. The solution is to determine user emotion by analyzing the rhythm of their typing patterns on a standard keyboard. A field study was conducted by collecting participants' keystrokes and their emotional states via self-reports. Keystroke features were extracted and classifiers for 15 emotional states were created by Epp, Clayton, Michael Lippold, and Regan L. Mandryk (2011) in [4].

Computers are acquiring the ability to express and recognize affect, and may soon be given the ability to have emotions. The essential role of emotion in both human cognition and perception, as demonstrated by recent neurological studies, indicates that affective computers should not only provide better performance in assisting humans, but also might enhance computers' abilities to make decisions. In [5] Picard and Rosalind W (1995) discusses key issues in affective computing that relates to arise or influence emotions. Models are suggested for computer recognition of human emotion and new applications are presented for computer assisted learning, perceptual information retrieval, arts and entertainment, and human health and interaction. Affective computing coupled with new wearable computers will also provide the ability to gather new data necessary for advances in emotion and cognition theory. A preliminary description of a novel type of chat system that aims at realizing natural and social communication between distant communication partners was discussed by Ma and Chunling (2005) in [6]. The system is based on an Emotion Estimation module that assesses the affective content of textual messages. Avatars associated with chat partners act out the assessed emotions of messages through multiple modalities, including synthetic speech and affect-related gestures.

Computing and communication systems have improved our way of life, but have also contributed to an increased data exposure and consequently to identity theft. A possible way to overcome this issue is by the use of biometric technologies for user authentication. Among the possible technologies to be analyzed, this work focuses on keystroke dynamics, which attempts to recognize users by their typing rhythm. In order to guide future researches in this area, a systematic review on keystroke dynamics was conducted and presented by Pisani, Paulo Henrique, and Ana Carolina Lorena (2013) in [7]. The systematic review method adopts a rigorous procedure with the definition of a formal review protocol. Systematic reviews are not commonly used in artificial intelligence. The process involved in the review along with the results obtained identifies the state of the art of keystroke dynamics. A summary of main classifiers, performance measures, extracted features and benchmark datasets used in the area. The need to secure sensitive data and computer systems from intruders while allowing ease of access for authenticating the user is one of the main problems in computer security. Traditionally, passwords have been the usual method for controlling access to computer systems but this approach has many inherent flaws. Keystroke dynamics is a biometric technique to recognize and an analysis of his/her typing patterns. In [8] Jyotsna Gaikwad et.al (2016) experiment measured mean, standard deviation and median values of keystroke features such as latency, duration, digraph and their combinations and compare their performance. The latest trend in authenticating users is by using the potentiality of biometrics. Keystroke dynamics is a behavioral biometrics which captures the typing rhythms of users and then authenticates them based on the dynamics captured. A detailed study on the evaluation of keystroke dynamics as a measure of authentication is carried out.

The use of biometrics for authentication mechanisms is becoming more and more important for research as well as industry. Keystroke dynamics is a biometric authentication method that improves the security of password-based applications. The performance of biometric keystroke recognition is still an open research issue. In fact, the extracted features relevant to a personal way of typing become less representative over time.This can lead to a failure in the

biometric verification task. Because of the changes of such features, the representative model has always to be updated. Use of growing and sliding windows as template update methods based on a statistical classifier was done in [9] by Mhenni, Abir, et.al (2016) to demonstrate that user-specific thresholds, varying from an update session to another which allows reducing the error rates compared to the update with a fixed threshold. A novel technique to strengthen password authentication system by incorporating multiple keystroke dynamic information under a fusion framework was done in [10] by Teh, Pin Shen, et.al (2010). They capitalize four types of latency as keystroke feature and two methods to calculate the similarity scores between the two given latency. A two layer fusion approach is proposed to enhance the overall performance of the system to achieve near 1.401% Equal Error Rate (EER). They also introduce two additional modules to increase the flexibility of the proposed system. These modules aim to accommodate exceptional cases for instance, when a legitimate user is unable to provide his or her normal typing pattern due to reasons such as hand injury.

Research on keystroke dynamics biometrics has been increasing, especially in the last decade. The main motivation behind this effort is due to the fact that keystroke dynamics biometrics is economical and can be easily integrated into the existing computer security systems with minimal alteration and user intervention. Numerous studies have been conducted in terms of data acquisition devices, feature representations, classification methods, experimental protocols, and evaluations. However, an up-to-date extensive survey and evaluation is not yet available.

In [11] Teh, Pin Shen, et.al (2013). objective was to provide an insightful survey and comparison on keystroke dynamics biometrics research performed throughout the last three decades, as well as offering suggestions and possible future research directions. In [12] Pantel, Patrick et.al (2002) presents a clustering algorithm called CBC (Clustering By Committee) that automatically discovers word senses from text. It initially discovers a set of tight clusters called committees that are well scattered in the similarity space. The centroid of the members of a committee is used as the feature vector of the cluster. The authors proceed by assigning words to their most similar clusters. After assigning an element to a cluster, removal of overlapping features from the element was done. This allows CBC to discover the less frequent senses of a word and to avoid discovering duplicate senses. Each cluster that a word belongs to represents one of its senses. Also an evaluation methodology for automatically measuring the precision and recall of discovered senses was presented. A detailed explanation of key terms involved including Word Sense discovery, different clustering algorithm comparisons, evaluation techniques and Machine learning algorithms was made.

Many different requirements can be placed on intrusion detection systems. One such important requirement is that it be effective. The system should detect a substantial percentage of intrusions into the supervised system, while still keeping the false alarm rate at an acceptable level. The proposed work in [13] by Axelsson, Stefan (2000) aims to demonstrate that, for a reasonable set of assumptions, contrary to what has previously been thought. The false alarm rate is the limiting factor for the performance of the intrusion detection system. This is due to the base-rate fallacy phenomenon that in order to achieve substantial values of the Bayesian detection rate, P (IntrusionjAlarm), the system should achieve—a perhaps unattainably low—false alarm rate, on the order of 1 105 , or 1=100; 000 per event. Unlike other access control systems based on biometric features, keystroke analysis has not led to techniques providing an acceptable level of accuracy was discussed by Bergadano, Francesco et.al (2002) in [14]. The reason is probably the intrinsic variability of typing dynamics, versus other--very stable---biometric characteristics, such as face or fingerprint patterns. An original measure for keystroke dynamics that limits the instability of this biometric feature was presented. The approach was tested on 154 individuals, achieving a False Alarm Rate of about 4% and an Impostor Pass Rate of less than 0.01%. The claimed performance was reached using the same sampling text for all the individuals, allowing typing errors, without any specific tailoring of the authentication system with respect to the available set of typing samples and users. The samples collected were over a 28.8-Kbaud remote modem connection.

Now-a-days people are heavily dependent on computers to store and process important information. User authentication and identification has become one of the most important and challenging issue in order to secure them from intruders. As traditional user ID and password scheme have failed to provide information security, keystroke dynamics authentication systems can be used to strengthen the existing security techniques. Keystroke dynamic authentication systems are transparent, low cost, and non-invasive for the user, but it has lower accuracy and lower performance compared to other biometric authentication systems. A detailed survey of the researches on keystroke dynamic authentication that have used neural networks for classification described in the last two decades was depicted in [15] by Brown, Marcus, and Samuel Joe Rogers (1993). The summary, accuracy of each experiment, and shortcomings of the researches have been presented in this study. Finally, the study addresses some challenges in keystroke dynamic authentication systems using neural networks that need to be resolving in order to get better performance. A method of providing security to keyboard based systems, by recognizing patterns of typing by a subject for identity confirmation, comprising the steps of, defining at least one statistical

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

relevance criterion that will qualify certain keystrokes in a group of keystrokes typed by a subject as a mini-rhythm, defining at least one enrolment phase criterion to indicate when text entered in an enrolment phase qualifies as meeting enrolment phase requirements: requiring a subject to enter an enrolment phase, analysing said plurality of sample text keystroke characteristic data against said statistical relevance criteria to identify if one or more groupings of sample text keystroke actions qualifies as a mini-rhythm and selectively using only mini-rhythm data from said sample text in [16] by Bender, Steven S., and Howard J. Postley(2007).

A method and apparatus is disclosed for verifying whether a particular individual is a member of a predetermined group of authorized individuals in [17] by Garcia and John D(1986). The subject apparatus is particularly suited for controlling access to a secure resource such as a computer network or data base. In accordance with the subject invention, time delays are measured between successive strokes of a keyboard as the individual enters his name. A timing vector, which is constructed from the time delays, is statistically compared with a stored timing vector derived from the authorized individual. If the timing vectors are statistically similar, the individual will be permitted access to the resource. Kevin S et.al [18] presents an anomaly detector for keystroke dynamics authentication, based on a statistical measure of proximity, evaluated through the empirical study of an independent benchmark of keystroke data. A password typing-rhythm classifier is presented, to be used as an anomaly detector in the authentication process of genuine users and impostors. The proposed user authentication method involves two phases. First a training phase in which a user typing profile is created through repeated entry of password. In the testing phase, the password typing rhythm of the user is compared with the stored typing profile, to determine whether it is a genuine user or an impostor. The typing rhythm is obtained through keystroke timings of key-down / key-up of individual keys and the latency between keys. The training data is stored as a typing profile, consisting of a vector of median values of elements of the feature set, and as a vector of standard deviations for the same elements. ROCR is a package for evaluating and visualizing the performance of scoring classifiers in the statistical language R. It features over 25 performance measures that can be freely combined to create two-dimensional performance curves. Standard methods for investigating trade-offs between specific performance measures are available within a uniform framework, including receiver operating characteristic (ROC) graphs, precision/recall plots, lift charts and cost curves. ROCR integrates tightly with R's powerful graphics capabilities, thus allowing for highly adjustable plots. Being equipped with only three commands and reasonable default values for optional parameters, ROCR combines flexibility with ease of usage in [19] by Sing and Tobias (2005).

III. PROPOSED SYSTEM

The automated process of detecting and authenticating a user on the rhythm of typing is referred to as keystroke dynamics or typing dynamics. Keystroke Dynamics falls under the class of behavioral biometric. Typed key measurements available from most every keyboard can be recorded to determine 'Dwell time' (The time a key pressed) and 'Flight time' (The time between "key down" and the next "key down"). The recorded keystroke timing data is processed through a unique neural algorithm, which determines a primary pattern for the future comparisons. The neural algorithm is provided with Digraph latencies (The elapsed time between the release of the first key and the depression of the second key). The extractions of such features are accepted from the free text provided for the user to create his own profile. Since keystroke dynamics works solely on typing rhythm, a user can be identified even if he uses a language different than the one he used initially. 'Intrusion detection system' has an accuracy of limiting the 'False alarms'. The system is very robust and will not authenticate intruders even if they use typing rhythms of a language that is different. By this we can say that the keystroke analysis is the next best solution to authenticate personal identity easily. A custom-built feature extractor was used in the pipeline for extracting 18 features from the key logger data. The features represent the characteristics of a user's typing pattern. The keyboard was divided into two parts as per the convention. The extracted features can be classified into four different categories such as Latency: It is the time interval between "KEY DOWN" event (pressing) of two consecutive key strokes. Hold Time: It is the time measure for which a particular key is pressed. Counts per Character: It is a fraction of the number of times a key is pressed over the total number of keys pressed in a keystroke data. Characters per Minute: It is the average number of keys pressed in a time interval of one minute. Data Acquisition, Keyword Extraction, Keyword Engineering and Classify acquired data are the functional requirements of the proposed system.

A. System Architecture for the proposed system

Figure 2 shows the system architecture of the proposed system. Features are calculated for entered text. Classifiers lies between the MaxMin threshold. Finally the authentication process to prove the legitimate user.

B. The requirements for the proposed system

Performance: The program should be highly portable so as to move it across computers. It is assumed that network connection will be available on the computer on which the program resides. Capacity, scalability and availability are the parameters considered to test the performance.

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

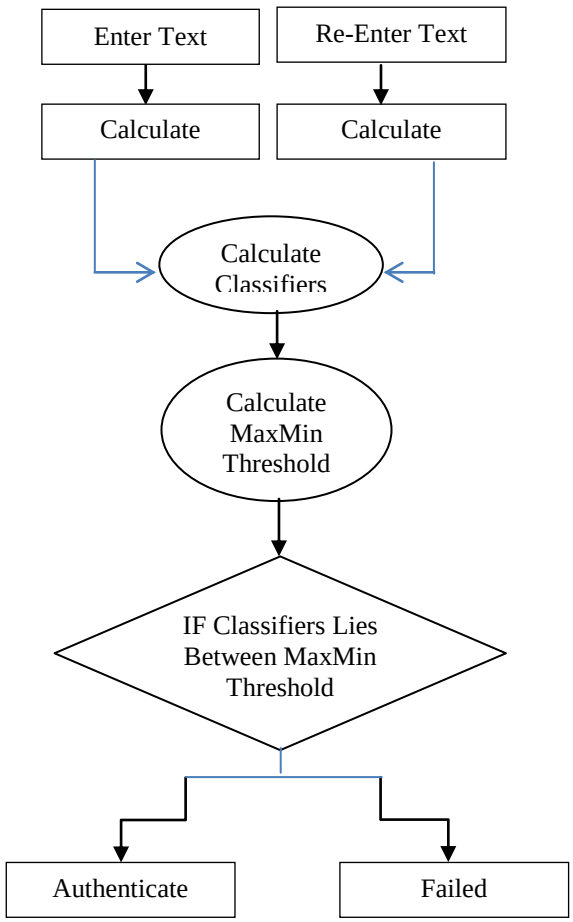

Fig. 2 System Architecture

- Safety: The data extracted for authentication cannot be used in any way that might harm the users or system.
- Security: Certifications are mandate as the system is used for authentication.
- Software Quality Attributes like maintainability, security, randomness, verifiability and load are considered for the proposed system.

Hardware: Core i5 with 2 GHz processor speed, 8 GB ram, 1TB hard disk and keyboard are used.

Software: Windows 7 or higher version of Operating system supports the implementation. Python2.7.14 is used as the coding language.

Software Interfaces: PyCharm as IDE, PyHook, Pythoncom, MatplotLib

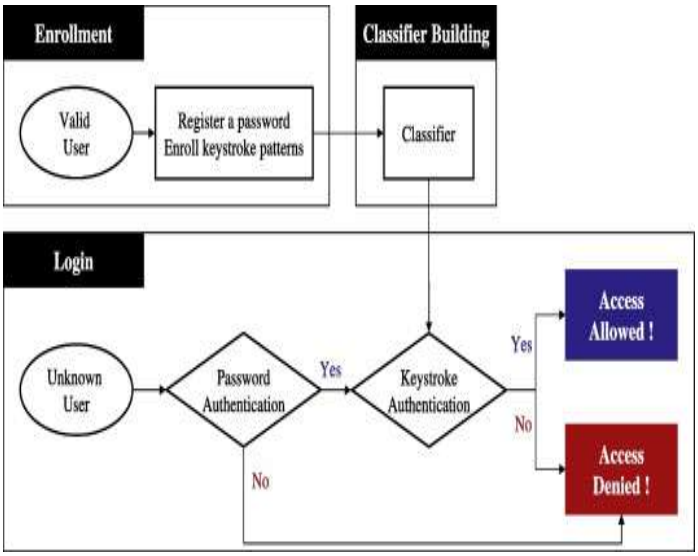

Fig. 3 Enrollment and Login

Enrolment and Login: The enrolment process is done using keystroke patterns of the valid user. The login process is done by the user. Valid user is authenticated using password and keystroke patterns compared with profiles stored in the database. Figure 3 represents the enrolment and login procedures.

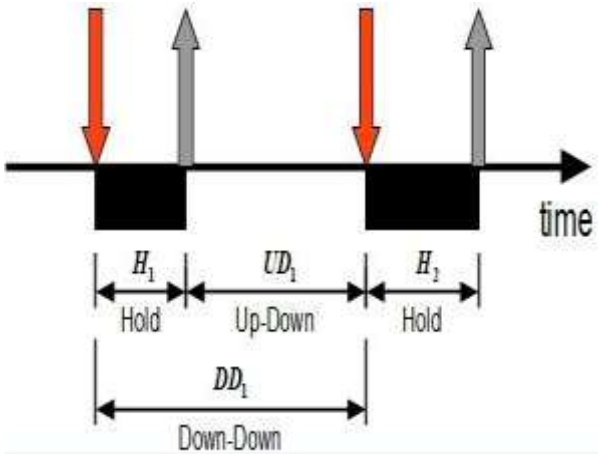

Fig. 4 Key Latency Transitions

Figure 4 shows the key latency transitions with annotations H,UD and DD.

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

Structure Chart: Logical decision to accept and reject a user is done using ANN, ANFIS and SVM. Figure 5 represents the structural chart of keystroke dynamics.

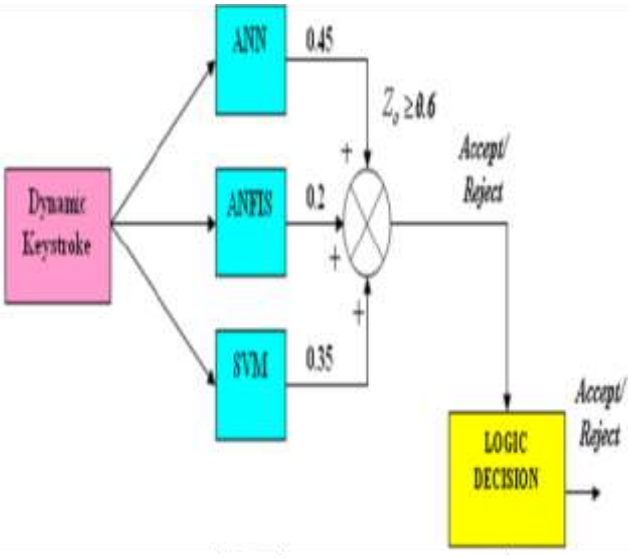

Fig. 5 Structure Chart of keystroke dynamics

Functional Description of Modules for enrollment and verification are shown in figure 6 and 7.

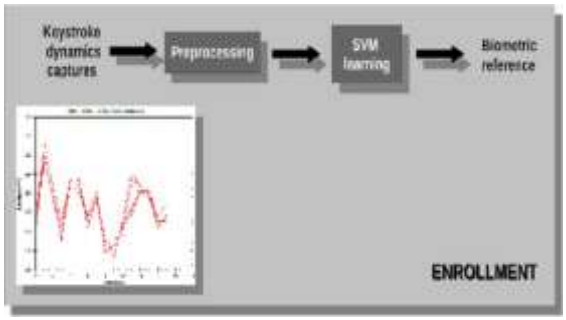

Fig. 6 Functional Description of Enrollment

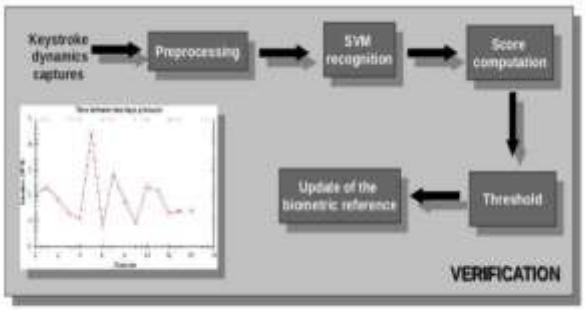

Fig. 7 Functional Description of Verification

IV. IMPLEMENTATION

A. Technologies

PyHook: The PyHook library wraps the low-level mouse and keyboard hooks in the Windows Hooking API for use in Python applications. Keyboard hooks work in the same manner as mouse hooks, but return different information.

Matplotlib: Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. We can generate plots, histograms, power spectra, bar charts, error charts, scatterplots, etc., with just a few lines of code.

Neural Networks: An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process.

Keystroke Dynamics: The automated process of detecting and authenticating an user on the rhythm of typing is referred to as keystroke dynamics or typing dynamics. Keystroke Dynamics falls under the class of behavioral biometric.

Since keystroke dynamics works solely on typing rhythm, a user can be identified even if he uses a language different than the one he used initially. The raw measurements used for keystroke dynamics are dwell time and flight time.

- ***Dwell time*** is the measure of how long a key is pressed
- ***Flight time*** is the interval of time taken between releasing a key and pressing the successive key.

B. Programming Language Selection and Code Conventions

Python: Python comes with a huge amount of inbuilt libraries. Many of the libraries are for Artificial Intelligence and Machine Learning. Some of the libraries are Tensorflow (which is high-level neural network library), scikit-learn (for data mining, data analysis and machine learning), pylearn2 (more flexible than scikit-learn), etc. Python has an easy implementation for OpenCV.
Python is an Interpreted language which in lay man's terms means that it does not need to be compiled into machine language instruction before execution and can be used by the developer directly to run the program. This makes it comprehensive enough for the language to be interpreted by an emulator or a virtual machine on top of the native machine language which is what the hardware understands.

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

C. Implementation Details of Modules

Data Acquisition Module: A key logger is used to collect the keystroke data for all the group members. This was collected in the form of Key Event - whether a button was pressed or released, Key Code - which button was pressed, Shift, Alt, Control - if these buttons were pressed or not, Timestamp - the exact timestamp of the event.

Feature Extraction Module: A custom-built feature extractor was used in the pipeline for extracting a set of 18 features from the key logger data. These features are characteristic of a user's typing pattern. The keyboard was divided into two parts as per the convention as shown in figure 8.

Fig. 8 Keyboard division

The extracted features can be classified into four different categories namely Latency: It is the time interval between "KEY DOWN" event (pressing) of two consecutive key strokes.

Hold Time: It is the time measure for which a particular key is pressed.

Counts per Character: It is a fraction of the number of times a key is pressed over the total number of keys pressed in a keystroke data.

Characters per Minute: It is the average number of keys pressed in a time interval of one minute.

The difference between timestamps aided in providing latency and hold times that were used in the training data. The processed data was further used to create bins of typing sessions of the user. Each bin represents a typing session of the user. A bin is created when there is no "Key Event" recorded for a period of ten seconds or when the user types uninterruptedly for a period of one minute without a break lasting more than ten seconds. Each bin was analysed for the latency and holds times for certain combination of keys. The latency and hold times have a threshold of 1.5 seconds above which they are ignored. The mean values of the latency and hold times of a bin for the combination of keys are considered as an observation vector.

A bin is considered as a valid bin if it contains more than hundred events within it, barring which it is discarded and the next bin is constructed. Additional features considered were the no. of backspaces per character of the user. This gives a sense of how prone to typing mistakes a user is and reflects mood of the user, although testing for this feature gives inconsistent results as this feature may also reflect the current nature of the text being typed and varied widely for the same user. Another feature used is the cpm(characters per minute), which reflects the typing speed of a user and which gives a distinction among several users. The cpm feature was considered in only those bins where the user typed more than 50 valid characters so as to give a true sense of a user's typing speed. Table 1 shows the extracted features.

Feature Engineering Module: First, the unnecessary keys captured by the Key logger, such as keys used for navigation, were filtered out as they did not have any significance in determining an user's typing pattern. Any missing values of an attribute for an observation bin are filled in by the mean of the corresponding features from the other bins. (*Filling in the missing values with the median results in a lower training and test accuracy) The data collected in this format gives us an observation vector wherein the user continuously types a stream of text, as opposed to the method of collecting latency data of character combinations at different time stamps and including them in a single observation vector.

Table 1: Extracted Features of key logged data

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

Latency	lr	Transition from left hand to right hand, without shift
	rl	Transition from right hand to left hand, without shift
	Lr	Transition from left hand with shift to right hand without shift
	Rl	Transition from right hand with shift to left hand without shift
	Ll	Transition from left hand with shift to left hand without shift
	Rr	Transition from right hand with shift to right hand without shift
	ll	Transition from left hand,to left hand without shift
	rr	Transition from right hand,to right hand without shift
Hold Time	l	Hold time of keys at left part without shift
	r	Hold time of keys at right part without shift
	L	Hold time of keys at left part with shift
	R	Hold time of keys at right part with shift
	SPACE	Hold time of space key
	ENTER	Hold time of Enter key
Characters per minute	CPM	Average number of keys pressed by the user in one minute

Classifier Module:

Mixture of Gaussians model (GMM) and 4 other different learning models was used for classifying the obtained training data from the feature extraction step. The GMM Estimator provided by the scikit-learn library in Python was used in this project. The means of all the attributes per output label class, was provided as an initialization parameter for the model. This mean, along with the co-variances and the weights, were improved over subsequent iterations. Evaluation was done using 4 different types of co-variances - Spherical, Diagonal, Full and Tied.

From the 500 vectors that every user is provided we use 200 sets to train the model and obtain the typing behavior of the user. The trained variable (a panda data frame) contains the training data.

Dataset Description:

For the training and testing datasets we have used Carnegie Mellon University's Keystroke dataset as published by them on :http://www.cs.cmu.edu/~keystroke/. It includes the rhythm pattern data information for 51 users, where each user has iteratively typed "tie5Roanl" 400 times. To further measure deviations; if any, a gap of one day was provided between different sessions to capture the user's data. Also for real time testing, we have also used data extracted with the help of the key logger.

V. EXPERIMENTAL RESULTS

System Testing is done for five learning models namely:

- Manhattan Distance
- Manhattan Filtered Distance
- Manhattan Scaled Distance
- Gaussian Mixture Model (GMM)
- Support Vector Machines (SVM)

Our first detector is Manhattan Detector (MD). The training() function calculates mean_vector for each user from the samples in train(training set). Only the mean of feature vectors is considered as user model. We now supply the model of a user with unseen test sample as explained here- a test_genuine list which has the remaining 200 feature vectors (repetitions of password) of the same user and a test_imposter list which has 5 vectors each from all the other 50 users making a total of 250 imposter samples. Each user will be tested 450 times for user's authenticity.
The detector calculates the city block distance (Manhattan distance) between the test samples and mean_vector for the subject. It is easy to understand that smaller values of this distance indicate higher similarity of the sample to the subject's model; however if the score has a larger value, it means the sample is quite dissimilar to the model and should get not get verified as the subject.

The obtained scores of genuine and imposter samples lists user_scores and imposter_scores respectively. We evaluate the equal error rate (EER) for the detector. The resulting Equal Error rate is presented as performance metric for each detector as-Average EER for Manhattan detector: 0.18065830919103248

The accuracy in predictions by the Mixture of Gaussian models has been shown in Table 2. It is observed that the Diagonal and Tied type Covariance Matrix performs better than the other types. For the test datasets, 4 out of 5 users were predicted with an accuracy of over 70%. Each individual test datasets were of the length of a standard news article.

Table 2 :Accuracy of Gaussian Models

Dataset	Tied(%)	Diagnol(%)	Spherical(%)	Full (%)
Training	86.29	79.44	43.15	42.83
User-1	70.37	85.19	92.59	88.89
User-2	94.74	94.74	78.95	100.00
User-3	77.78	100.00	11.11	-
User-4	89.29	53.57	25.00	-
User-4	60.00	60.00	-	-

Table 3: Trained Values of classifier

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

Dataset	1	2	3	4
Training	87.65	76.29	89.41	90.27
User-1	-	70.37	88.89	-
User-2	-	94.74	100	100.00
User-3	-	77.78	-	22.22
User-4	7.14	89.29	92.86	92.86
User-4	20.00	60.00	20.00	-

Table 3 depicts the final model trained used mean to fill up missing values in bins, unused number of backspace per character as a feature and considered bins with more than 100 events and latency and hold time thresholds of 1.5 s.

Table 4 consolidates all the Equal Error Rates of the detectors and tare used for table 5 shows the comparison of performance metrics for the proposed system.

Table. 4: EER Values of detectors

Detector	Average Equal Error Rate
Manhattan Detector	0.18
Manhattan Filtered Detector	0.15
Manhattan Scaled Detector	0.12
One Class SVM Detector	0.12
GMM Detector	0.15

Table 5: Comparison of Performance Metrics

Detector	Average Equal-Error Rate	Standard deviation of EER
Manhattan Scaled Detector	0.0945	0.068375
Outlier Count (z-score)	0.103167	0.07691
Nearest Neighbor (Mahalanobis)	0.1075	0.06213
SVM (one-class)	0.12068	0.0586
Manhattan Filtered	0.12535	0.081299
Mahalanobis	0.1337	0.06678
Mahalanobis Normed	0.1337	0.06678
Manhattan	0.15	0.09
K-Means	0.1559	0.072
Neural Network (auto-assoc)	0.16417	0.0914199
Euclidean	0.16929	0.0931429
Euclidean Normed	0.2107	0.1174
Neural Network (standard)	0.6551	0.1866

Key Findings

Testing the system using multiple detectors we found that MSD and SVM had the lowest average. EER and MD had the highest.
We also found that though physiological biometrics is more reliable than behavioral biometrics. The criteria to measure the suitability of keystroke dynamics are:

- *Universality:* Any keyboard user can take advantage of this solution.
- *Uniqueness:* Behavioral biometric differs from physiological biometrics in the sense that there might not be something like an absolute match. This might cause difficulty to ascertain the uniqueness of the rhythm. The FRR and Far levels of keystroke dynamics will not be as good as those obtained by good physiological biometric factors, though these are not the only factors that determine the level of authenticity.
- *Performance:* Different emotional and physical states of the user may make the user's typing rhythm vary during the same day. This is cause for a major problem with using keystroke analysis
- *Collectability:* Collectability refers mainly to the cost estimated with the use of the product. The major advantage of keystroke dynamics is that external hardware which might cost more is not needed. It also

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

possible obtain and store keystroke patterns in the background without causing much or no overhead.

- Acceptability: Depending of the country or state you are in using key logging software might be a direct violation of local laws. Even if the actual typed text is not analyzed or retained, applicable legislation is sufficiently unclear to be in your disadvantage when you intend to actually use keystroke dynamics. Request legal advice before implementing or experimenting without written consent from people on the keyboard.
- Circumvention: It is certainly difficult, if not impossible to mimic another person's typing rhythm. Electronically capturing using keylogging software is possible, thus implementing this biometric solution requires that data security is guaranteed from the input (keyboard) to the matching algorithm.
- Performance: Behavioral biometrics have higher variations because they depend on a lot of (external) factors such as ergonomics, fatigue, mood, etc. This causes higher FAR and FRR when compared to solutions based on a physiological biometric factor such as fingerprint recognition.

CMU Keystroke Dynamics Benchmark Data set is used for testing. Data extraction and classification are done according to the table mentioned. Sample of the data extracted as a csv file and stored according to the given key. Calculation of latency of time, hold times and no. of backspaces and cpm are done. Coding is done to store features required and remove unwanted data.

Manhattan Filtered Detector: The training() function of the MFD takes into account any outliers in a subject's typing habits. Such deviations from his/her usual typing habits may occur due to a variety of reasons, like the user being tired or bored and hence typing exceptionally slower than normal, etc. MFD simply filters(removes) such outliers and calculate mean_vector and standard deviation vector, std_vector for the user from his training vectors.

To reject the outliers, first the euclidean distance between each of the training vectors and mean_vector is determined. Then, any vector for whom this distance is greater than three times the std_vector is an outlier and is dropped from train. Dropping_indices consists of the indices of all such outliers. Having eliminated all such tarining vectors from the set, mean_vector for the user is re-calculated from the remianing samples in train. This mean_vector is now the model for the user's typing behavior.

Averegage EER for Manhattan Filetered detector: 0.148487121188

Manhattan Scaled Detector: While training, we calculate the mean_vector() as well as the mad_vactor() which has the mean absolute deviation(MAD) of each featre of the training data. In testing(), score for a test sample is being calcuated as where xi and yi are the i^{th} feature in the test sample and mean_vector() respectively, and αi is that feature's MAD, taken from the mad_vector(). Hence we essentially calculate the city-block distance but each feature is getting scaled by its MAD.

The resulting average EER for Manhattan scled detector : 0.117636962313

Thus, in the detectors based on the Manhattan distance as the similarity score, we see that $EER_{MSD} < EER_{MFD} < EER_{MD}$, making MSD the superior one.

One class SVM: One class SVMs learn a decision function from the data of one class only and test a new sample to found out whether it is like the training data or not. We use sklearn.svm.OneClassSVM sub-module's fit() function to train a OneClassSVM object, named clf here, and the decision_function() function to calculate the similarities scores for the test samples.
The Average EER of one-class SVM detector is: 0.12065079948315142

Limitations:

Keystroke Dynamics is not as robust as other physiological biometric methods such as Fingerprint analysis or iris recognition. Also keyboards in different countries like Europe and China might be in different languages. Figure 9 shows the comparison of biometric features.

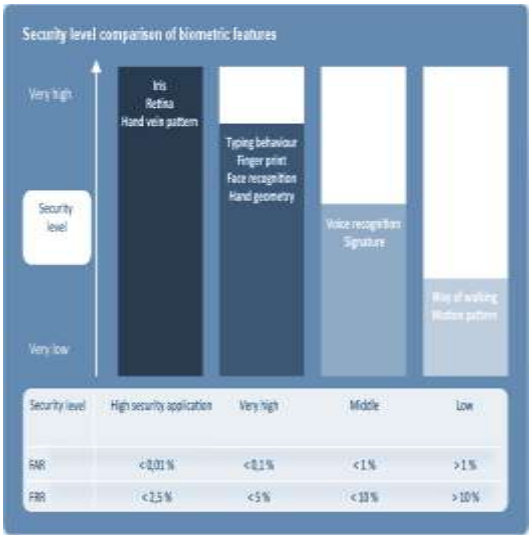

Fig. 9 Comparison of Biometric Features

International Journal of Engineering Applied Sciences and Technology, 2019
Vol. 4, Issue 3, ISSN No. 2455-2143, Pages 213-225
Published Online July 2019 in IJEAST (http://www.ijeast.com)

Difficulties encountered and Strategies Applied

- Data extracted through key logger did not have all the information required to test the various classifiers.
- So the dataset provided by Carnegie Mellon University of the US was used to test.

VI. CONCLUSION & FUTURE ENHANCEMENT

We show the uniqueness and usefulness of keystroke dynamics as an authentication measure using the typing rhythm of the user. Keystroke dynamics can be used for long samples of text which further can be used to increase the robustness of the method.

Though Keystroke dynamics has been discovered since WW II, a measure to incorporate it in daily authentication systems have not been made. Knowing the cost effectiveness and efficiency we discovered when we tested the algorithm, we find that a much better use of keystroke analysis for authentication need to be used. We could introduce keystroke dynamics based authentication and present distance metric or scoring based machine learning models. We finally believe that keystroke dynamics is an underappreciated system and can be used extensively as a robust authentication tool either individually or as a hybrid system.

Keystroke Dynamics can be used as a very specific form of surveillance**.** There can exist software solutions which, often with end-users being aware of it, track keystroke dynamics for each user account. This tracking, historization of keystroke dynamics is then used to analyze whether accounts are being shared or in general are used by people different from the genuine account owner. Reasons for such an implementation could be verification of users following security procedures and to verify that no software licenses are being shared especially for SAAS applications. Test various other detectors, like neural networks, Mahalanobis distance and compare their performance. Find ways to capture their own keystroke information like one used in database and verify themselves with the system.

VII. ACKNOWLEDGMENT

The authors would like to acknowledge BMS college of Engineering and TEQIP III phase for their immense support in carrying out and encouraging this research work.

VIII. REFERENCE

[1] Chang and J. Morris (2013) "Capturing cognitive fingerprints from keystroke dynamics." IT Professional 15.4 (pg 24-28).

[2] Fan and Xian (2016) "Apply word vectors for sentiment analysis of app reviews." 2016 3rd International Conference on Systems and Informatics (ICSAI). IEEE.

[3] Mondal, Soumik, and Patrick Bours (2017). "Person identification by keystroke dynamics using pairwise user coupling." IEEE Transactions on Information Forensics and Security 12.6 (pg 1319-1329).

[4] Epp, Clayton, Michael Lippold, and Regan L. Mandryk (2011) "Identifying emotional states using keystroke dynamics." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM.

[5] Picard and Rosalind W (1995) "Affective Computing-MIT Media Laboratory Perceptual Computing Section Technical Report No. 321." Cambridge, MA 2139.

[6] Ma and Chunling (2005) "A chat system based on emotion estimation from text and embodied conversational messengers." Proceedings of the 2005 International Conference on Active Media Technology, (AMT) IEEE.

[7] Pisani, Paulo Henrique, and Ana Carolina Lorena (2013) "A systematic review on key stroke dynamics." Journal of the Brazilian Computer Society 19.4 (pg 573).

[8] Gaikwad and Jyotsna (2016) "User Authentication using Keystroke Dynamics."

[9] Mhenni and Abir (2016) "Keystroke template update with adapted thresholds." 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). IEEE.

[10] The and Pin Shen (2010) "Keystroke dynamics in password authentication enhancement." Expert Systems with Applications 37.12 (pg 8618-8627).

[11] Teh, Pin Shen, Andrew Beng Jin Teoh, and Shigang Yue (2013) "A survey of key stroke dynamics biometrics." The Scientific World Journal 2013.

[12] Pantel, Patrick, and Dekang Lin (2002) "Discovering word senses from text." Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM.

[13] Axelsson and Stefan(2000) "The base-rate fallacy and the difficulty of intrusion detection." ACM Transactions on Information and System Security (TISSEC) 3.3 (pg 186-205).

[14] Bergadano, Francesco, Daniele Gunetti, and Claudia Picardi (2002) "User authentication through keystroke dynamics." ACM Transactions on Information and System Security (TISSEC)5.4 (pg 367-397).

[15] Brown, Marcus, and Samuel Joe Rogers(1993) "User identification via keystroke characteristics of typed names using neural networks." International Journal of Man-Machine Studies 39.6 (pg 999-1014).

[16] Bender, Steven S., and Howard J. Postley (2007) "Key sequence rhythm recognition system and method." U.S. Patent No. 7,206,938.

[17] Garcia, John D (1986) "Personal identification apparatus." U.S. Patent No. 4,621,334.

[18] Killourhy, Kevin S., and Roy A. Maxion (2009) "Comparing anomaly-detection algorithms for keystroke dynamics." 2009 IEEE/IFIP International Conference on Dependable Systems & Networks. IEEE.

[19] Sing and Tobias (2005) "ROCR: visualizing classifier performance in R." Bioinformatics 21.20 (pg 3940-3941).